où l'Insee expérimentait la nouvelle nomenclature des catégories sociales (CS), on a jugé de circonstance d'opérer une double codification des professions selon la nomenclature des catégories socioprofessionnelles (CSP) de 1954 et des professions et catégories socioprofessionnelles (PCS) de 1982. Enfin, un programme de codification automatique des communes a été mis au point, compte tenu des 150 000 mentions dans les questionnaires. L'indication sur écran des numéros de département et des toutes premières lettres de l'intitulé de la commune suffisant à son identification et à sa codification.

4.2. Fichiers de base de l'enquête

Il existe une base SAS de l'enquête qui est constituée d'un enregistrement par individu. Une partie des questions n'y figure cependant pas.

4.3. Choix de la saisie de la variable temps

La saisie des mois et années déclarées dans le questionnaire a conduit à une élaboration originale. Le traitement des dates mal déclarées a retenu une formule de correction sans falsification des dates déficientes : le codeur substituait à une date absente ou incohérente la date centrale des intervalles disponibles en en portant la mention sur le code du mois de l'événement. Les valeurs 21 à 32 indiquaient l'omission du mois, 41 à 52 celle de l'année, 61 à 72 signifiaient que la date originelle était incompatible avec le reste du questionnaire.

4.4. Choix de l'unité de temps pour l'analyse

Si l'on a parfois utilisé le mois, la plupart des analyses n'ont tenu compte que de l'année de l'événement.

4.5. Utilisation des méthodes d'analyse des biographies

Les logiciels suivants ont été utilisés :

— ROOT.RAT et EVACOV.FOR (écrits par Éva Lelièvre) ;
— RATE (écrit par Nancy Tuma) ;
— SAS ;
— TDA (écrit par Goetz Rohwer) ;
— STATA (routines reprenant les programmes ROOT.RAT et EVACOV.FOR écrites par Philippe Bocquier).

4.6. Autres méthodes d'analyse

Tris simples ou multiples, régressions, modèles logit, modèles multi-niveaux.

4.7. Les utilisateurs des données

— Brigitte Baccaïni (thèse, articles) ;
— Marie-Ange Cambois (article) ;

en 9 postes fondée sur type de métier et qualification[18]. Nous avons aussi constitué une variable agrégée en 4 postes, qui est, cette fois, une échelle. Le premier niveau regroupe personnels de service, employés et ouvriers non qualifiés (la mobilité à l'intérieur de cette catégorie a été fréquente au cours de la vie de travail), le second regroupe employés et ouvriers qualifiés.

4.2. Fichiers de base des suivis

Nous avons évoqué plus haut la complexité du corpus des données (cf. 1.3). Nous avons rassemblé autour d'un fichier principal une vingtaine de fichiers correspondant aux différentes sources documentaires (3 enquêtes principales, actes de naissance et de décès...) et aux différentes unités d'observation (enquêtés, résidences, états matrimoniaux, enfants...).

Les histoires matrimoniale, professionnelle et résidentielle sont saisies dans autant de fichiers différents. L'histoire résidentielle est répartie en 2 fichiers : un fichier « mort » correspondant à l'histoire rétrospective antérieure à l'enquête de 1985 et un fichier « vivant » correspondant à l'histoire post-retraite et mis à jour annuellement. Il faut ajouter qu'un certain nombre de dates remarquables dans la vie de l'enquêté se trouvent dans d'autres fichiers : arrivée à Paris, acquisition d'une résidence secondaire, début de cohabitation avec de vieux parents, etc.

4.3. Choix de la saisie de la variable temps

Jusqu'à présent les dates ont été saisies avec une précision à l'année, exception faite de la naissance connue au jour près. Compte tenu de la durée des séjours en institution, qui peuvent être très courts, les dates d'entrée en institution et de décès sont saisies au jour près.

4.4. Choix de l'unité de temps pour l'analyse

Soit nous utilisons des durées annuelles (définies par différences de dates annuelles), soit nous employons des couples avant/après un événement choisi. Pour l'âge au décès ou les durées de séjours finaux en institution, nous pouvons travailler au jour près.

4.5. Utilisation des méthodes d'analyse des biographies

Nous n'avons pas eu recours à de telles méthodes. Ce n'est pas une position de principe et nous comptons le faire. Mais nous savons que notre corpus n'est pas achevé : un cinquième de nos enquêtés sont encore vivants et nous n'avons pas épuisé les possibilités de mises à jour de nos données.

4.6. Autres méthodes d'analyse

Nous avons fait, bien sûr, de très nombreuses tabulations, et quelques analyses factorielles. Nous avons très souvent constitué des listes d'enquêtés selon des critères choisis afin de retourner aux documents matériels, de les

D'une manière générale, on a sous-estimé la mobilité du parc lui-même. Sans changer de logement, l'enquêté pouvait vivre dans un appartement dont le nombre de pièces s'était modifié ; son statut d'occupation pouvait également changer. Les cas les plus fréquents étaient le passage du statut d'hébergé par la famille à celui de locataire ou propriétaire (les parents décèdent et l'enfant reprend le bail à son nom ou encore hérite du logement, le passage de locataire d'un logement loi de 1948 à propriétaire par exemple).

4.2. Fichiers de base de l'enquête

Plusieurs fichiers existent :

— un fichier contenant les principales variables démographiques et professionnelles ;

— un fichier logement avec toutes les données de chaque période résidentielle ;

— un fichier corésidents ;

— un fichier sur les différents membres de la famille (lieu de résidence, statut d'occupation...) ;

— un fichier sur les autres logements, les résidences secondaires, les projets de retraite et le logement idéal.

4.3. Choix de la saisie de la variable temps

La variable temps a été saisie à partir du mois et de l'année.

4.4. Choix de l'unité de temps pour l'analyse

C'est l'année qui a été le plus souvent retenue. Les mois ont été pris en compte dans le calcul de durée ou d'âge.

4.5. Utilisation des méthodes d'analyse des biographies

ROOT.SAS (mis au point par Éva Lelièvre)

Autres Logiciels utilisés pour l'exploitation :

— logiciels de saisie de l'Ined : SID ;

— logiciels de traitement statistique : SAS, SPAD, SPADT.

4.6. Autres méthodes d'analyse

Analyses sur les trajectoires familiales, résidentielles et géographiques.

4.7. Les utilisateurs des données

Des chercheurs essentiellement :

— Catherine Bonvalet, Ined ;

— Dominique Maison, Ined, Université Paris X ;

— Paul White, Professeur à l'Université de Sheffield, Grande-Bretagne ;

4. Traitement de l'information

4.1. Traitement du questionnaire, principes et difficultés de codification

Les premières données tirées de l'enquête ont été exploitées par l'Office central de statistique en collaboration étroite avec l'Institut de statistique et de démographie, à l'École d'économie de Varsovie. L'ensemble du procédé de codification et de contrôle a été exécuté par l'Office central de statistique. Afin de vérifier la cohérence des données du questionnaire, nous avons mis au point un programme spécial qui enregistre et contrôle la cohérence des informations. Le programme de base se composait de 200 sous-programmes chargés de s'assurer de la qualité des données.

Dans un deuxième temps, l'Institut de statistique et de démographie recevait les données de l'Office central de statistique. Au sein de l'Institut ont été élaborés de nouveaux programmes destinés à vérifier les données et s'appuyant sur des options spécifiques pour évaluer la qualité des données par rapport au procédé des échantillons et à l'étude de la cohérence des structures et répartitions empiriques en général.

4.2. Fichiers de base de l'enquête

L'enquête rétrospective polonaise de 1988 a un manuel d'utilisateur en version polonaise et anglaise. Les données sont disponibles aux formats TXT, DBF, SPSS et ZOO.ARCHIVE pour le programme TDA. Les données sont organisées suivant les dix sections du questionnaire.

4.3. Choix de la saisie de la variable temps

Les variables temps n'ont pas été tout de suite entrées dans la base de données. Les bases de données comprennent les dates des événements, par exemple : la date du premier mariage – le mois et l'année – et la date de la première dissolution. L'utilisateur des bases de données peut créer des variables temps, dans le cas présent la durée du mariage.

4.4. Choix de l'unité de temps pour l'analyse

Le temps, selon l'analyse, est envisagé comme :

— le temps calendaire ou l'âge,

— la durée du processus depuis son origine ou la durée de séjour dans l'état étudié.

4.5. Utilisation des méthodes d'analyse des biographies

De nombreuses méthodes de pointe d'analyse biographique ont été utilisées au cours du travail de recherche sur les données de l'enquête. Cela n'a posé aucun problème, au contraire nous pensons avoir contribué au développement de ces méthodes.

4. TRAITEMENT DE L'INFORMATION

4.1. Traitement du questionnaire, principes et difficultés de codification

Dans le questionnaire, il n'y avait pas de questions ouvertes et cela a simplifié nettement la codification. Quelques problèmes se sont posés pour l'identification de la réponse à certaines questions. Par exemple, le type de localité (il n'a pas toujours été facile de délimiter le centre historique par rapport au reste de la ville, ou celui-ci de la zone suburbaine), le statut professionnel (la distinction entre employé et ouvrier est particulièrement floue), le type de logement (comment par exemple, classifier une maison de campagne rénovée qui ressemble à une petite villa indépendante).

Pour le reste, les codes étaient déjà spécifiés dans la case de réponse du questionnaire. La saisie a été effectuée en « couple » – l'enquêteur et un de mes collaborateurs – et à ce stade, cela a permis de réduire les erreurs à un niveau négligeable.

4.2. Fichiers de base de l'enquête

Le fichier de base, à partir duquel des fichiers allégés à format fixe ne contenant que des informations circonstanciées sont extraits lors des différents traitements, est un fichier formé d'un premier enregistrement consacré aux informations générales sur le ménage enquêté et aux données sur ses perspectives de logement. Il s'agit de toutes les informations contenues dans les volets 1 et 5 intégrées sous la forme de trois chiffres qui indiquent respectivement le nombre de membres du ménage, le nombre d'activités professionnelles enregistrées et celui des logements occupés. Pour chacune de ces séries d'informations, un enregistrement suit et donc : un pour chaque personne qui fait ou a fait partie du ménage, un pour chaque activité professionnelle exercée durant son appartenance au ménage, un pour chaque logement occupé. Chacun de ces quatre types d'enregistrement a son propre format fixe, mais le nombre d'enregistrements relatif à chaque ménage enquêté reste, bien sûr, variable.

4.3. Choix de la saisie de la variable temps

Je n'ai aucun regret d'avoir décidé de me contenter du niveau de précision de l'année, même si cela augmente les situations d'apparente contemporanéité des événements. Du reste, comme l'a déjà démontré J. Duchêne[1], quand on sollicite la mémoire à propos d'événements lointains, il est souvent difficile de se rappeler en quelle année un événement est survenu. De toute façon, la structure du questionnaire a empêché de vérifier, durant l'entretien, la coïncidence des dates pour en préciser éventuellement la succession.

[1] J. Duchêne, 1985, Un test de fiabilité des enquêtes rétrospectives « Biographie familiale, professionnelle et migratoire », Chaire Quetelet : *Migrations internes*, Louvain la Neuve.

Cette structure est reproduite par le programme de saisie : chaque questionnaire comprend autant d'enregistrements informatiques que d'enfants, de périodes d'emploi et de résidences principales.

Le codeur va substituer, à une année absente ou incohérente, l'année centrale de l'intervalle des possibles en en portant mention sur le code de l'année de l'événement. On a retenu cette formule de correction des dates déficientes, utilisée par l'Ined, car cet artifice de recodage est d'une grande efficacité informatique.

4.2. Fichiers de base de l'enquête

Rappelons que nous n'avons recueilli que les résidences ou les emplois occupés durant au moins six mois. Pour la cohérence des biographies, une succession de périodes de moins de six mois, totalisant cette durée, sera décrite comme une période unique désignée par « résidences variables », « types d'activité variables », « lieux d'emploi variables ». Cette consigne a cependant pour inconvénient d'éliminer la majorité des périodes de maladie et de chômage, et peut donner lieu à une variation par suite du mode de calcul.

Pour la saisie des questionnaires, nous avons utilisé le programme SID (Saisie interactive des données). À l'Ined, ce programme ne réalise que les contrôles de champ (chaque variable ne prend que les valeurs admises). La liste des codes admis est fermée. Le programme de saisie SID de notre enquête « 3B.R », tel que nous l'avons élaboré à l'Ined contient 27 blocs ; pour le bloc « enfants » nous avons prévu 15 occurrences et pour les blocs « emplois et logements », 30 occurrences. La saisie a été réalisée par le Centre d'informatique de la Commission nationale de statistique de la Roumanie, à la suite d'un accord passé avec l'Ined.

Les *questions ouvertes* seront traitées à part, par la méthode sociologique de l'analyse du contenu. On a trié un certain nombre de questionnaires, les mieux remplis, en espérant qu'ils illustrent de façon concrète une situation statistiquement fréquente. Les questions ouvertes ont été soumises à l'analyse de contenu, les catégories de l'analyse étant élaborées et établies sur la base de la lecture et de l'analyse de ces réponses. Nous avons assuré la relecture des questionnaires et l'analyse de contenu.

Les fichiers de base de l'enquête sont : la base d'analyse SAS et la base d'analyse par TDA.

4.3. Choix de la saisie de la variable temps

La saisie de la variable temps se fait par année.

4.4. Choix de l'unité de temps pour l'analyse

Dans l'analyse, nous avons traité le temps par des variables de durée. La particularité des données de durées est qu'elles peuvent s'interpréter facilement comme résultant d'un processus stochastique sous-jacent. Ce processus

mois, dans les cas de période de transition (par exemple, le chômage entre deux emplois), la date de fin d'une période correspond souvent à la date de début de la période suivante qui est généralement mieux connue. On a tenté, dans d'autres enquêtes, de réduire le problème en demandant, par exemple, le temps mis à trouver un emploi (voir plus haut).

4.3. Choix de la saisie de la variable temps

La saisie s'est faite au mois et à l'année avec code particulier pour situer l'ordre des événements simultanés (voir plus haut). Au moment de l'analyse, une datation au mois approximative est redonnée : 3e mois pour le code 33, 6e mois pour 66, 9e mois pour 88, et 6e mois pour 99 (dans le cas où il n'y a pas d'événements simultanés la même année).

À l'issue du terrain on se retrouve avec une série d'informations biographiques concernant un même individu. Le passage du questionnaire aux fichiers informatiques constitue un travail délicat. Dans le cas de l'enquête de Dakar, nous disposons en parallèle de l'itinéraire migratoire, de l'itinéraire professionnel à Dakar, de l'itinéraire résidentiel en ville, de l'itinéraire matrimonial, de la descendance. Pour chaque thème nous constituons un fichier où, pour un même individu, les événements sont datés et classés dans le temps, en ramenant toutes les dates à une durée écoulée en mois depuis le 1er janvier 1900. Ensuite il faut fusionner ce fichier avec les autres fichiers concernant les autres itinéraires, afin d'obtenir, dans un fichier unique, l'ensemble des événements vécus par l'individu, classés dans le temps.

Pour un même individu, le fichier comprend autant de lignes que de changements d'état, et ainsi on peut savoir à chaque instant de sa vie, sa profession, son lieu de résidence, sa situation matrimoniale, la naissance de ses enfants et les renseignements afférents à ces divers événements, et mettre ces éléments en relation. Ces différentes informations sont datées, et l'on connaît, pour chaque période de la vie de l'individu, le temps passé dans un état.

4.4. Choix de l'unité de temps pour l'analyse

L'unité de temps est l'année pour les événements simples, et le mois pour l'analyse des interférences avec d'autres événements (grâce au système de datation dans l'année). Dans la pratique, toute la datation se fait au mois près, car les données de chaque module (emploi, logement, mariages) ont été fusionnées pour qu'apparaisse clairement la succession des événements. Quand un seul événement a lieu pour une année, et si on ignore son mois d'occurrence, l'événement est arbitrairement daté au mois de juin de l'année.

Pour faire l'analyse du statut migratoire, nous avons constitué 3 groupes :

1. les enquêtés présents à Dakar à l'âge de 12 ans : certains parmi eux sont migrants (n'étant pas nés à Dakar), mais nous les considérons comme ayant été formés (à l'école ou en apprentissage) à Dakar : ils ont subi à un âge crucial les possibilités et les contraintes de la ville en matière de formation ;

4.3. Choix de la saisie de la variable temps

La variable de temps retenue permettait de situer chacun des événements en mois et en année. L'incapacité à produire une précision mensuelle pour certains événements nous a conduits indirectement à retenir une précision semestrielle ou trimestrielle (codes 33, 66, 88 et 99).

4.4. Choix de l'unité de temps pour l'analyse

La plupart des événements sont calculés en terme de durées écoulées en mois depuis le 1er janvier 1900.

4.5. Utilisation des méthodes d'analyse des biographies

Nous utilisons les méthodes d'analyse des biographies (descriptives et multivariées) qui se trouvent notamment dans le manuel d'analyse des biographies de Bocquier (1996).

4.6. Autres méthodes d'analyse

Aucune.

4.7. Les utilisateurs des données

Professeurs, chercheurs, doctorants, étudiants à la maîtrise, principalement de l'Université de Montréal et du Cerpod : Victor Piché, Dieudonné Ouedraogo, Richard Marcoux, Mamadou Kani Konaté, Mouhamadou Gueye, Michel Diawara, Amadou Bâ, Aka Kouamé, Richard Morin, Lucie Gingras, Thomas Legrand, Philippe Bocquier, Soumaila Mariko, Mireille Kantiebo, Fatouma Sissoko, Seydou Souley, Chantal Rondeau, Gaël LeJeune.

4.8. Confrontation avec d'autres enquêtes biographiques

Notamment avec les résultats de l'« Enquête qualitative sur l'appropriation foncière à Bamako » de 1993 (EQAF-Bamako, 1993). Ceci nous a permis de rendre compte des difficultés à cerner le statut d'occupation du logement à l'échelle des individus, des ménages et des familles (familles élargies). Nous avons également analysé certains de nos résultats (rétrospectifs) en les comparant avec les données des enquêtes et recensements réalisés au Mali depuis 1960. Par ailleurs, notre équipe travaille étroitement avec l'équipe de l'Ifan/Orstom qui a réalisé une enquête similaire dans la capitale sénégalaise en 1989. Nos travaux d'analyse comparée viennent de paraître[2].

[2] Voir Ph. Antoine, D. Ouedraogo, V. Piché (éds), 1998, *Trois générations de citadins au Sahel*, L'Harmattan, Collection Villes et entreprise, 276 p.

4.2. Fichiers de base de l'enquête

Le fichier de base de l'enquête est un fichier ASCII ayant un dictionnaire séparé. Comme nous l'avons déjà signalé, il a été produit à l'aide du logiciel ISSA, et a une structure arborescente. Il y a, d'une part, un fichier ménage et, d'autre, part six fichiers biographiques, un par module. Ces fichiers biographiques sont fusionnés ultérieurement pour constituer un fichier biographique unique où tous les événements sont classés selon leur date d'occurrence.

Les échéances des différentes trajectoires ont été saisies au mois près.

À l'issue de la saisie, les données ont été exportées par module vers le logiciel STATA. Avec le concours d'un expert en traitement des données biographiques, les fichiers ont été nettoyés à l'aide des programmes STATA. Les incohérences, la non déclaration des dates étaient imprimées sur listing avec l'identification du questionnaire biographique. Deux types de problèmes étaient courants :

— l'absence des dates dans le questionnaire bien que celles-ci figurent sur la fiche AGEVEN ;

— les événements (union, activité, résidence, etc.) n'étaient pas souvent enregistrés dans l'ordre chronologique dans le questionnaire.

En recourant aux questionnaires et aux fiches AGEVEN, certaines dates ont été retrouvées. Les événements pour lesquels les mois n'étaient pas déclarés étaient supposés avoir lieu au mois de juillet.

4.3. Choix de la saisie de la variable temps

La variable temps n'a pas subi de modification par rapport à la collecte, et l'unité temporelle de saisie reste le mois. Il faut juste signaler que le septième mois a été utilisé en remplacement des mois inconnus.

4.4. Choix de l'unité de temps pour l'analyse

Pour l'analyse, l'unité de temps utilisée est le mois. Pour chacun des événements pris en compte, le temps a été ramené à la durée écoulée en mois depuis le 1er janvier 1900, méthode qui s'avère particulièrement efficace pour l'analyse des biographies.

4.5. Utilisation des méthodes d'analyse des biographies

Un large recours aux techniques d'analyse des transitions (courbes de séjour de Kaplan-Meier, courbes à risques concurrents de Aalen, modèles semi-paramétriques de Cox) a été fait, et ceci sans grande difficulté. Cette facilité d'utilisation des méthodes d'analyse des biographies s'explique essentiellement par la convivialité du logiciel utilisé : le logiciel STATA qui, en plus, consomme peu de temps et d'espace.

de façon à réduire le travail de codification ou de transformation de l'information collectée et donc à minimiser le risque d'erreur, et, d'autre part, produire directement des fichiers ayant la structure adéquate pour leur traitement et leur analyse. Ce sont ces critères qui nous ont conduites à une organisation générale de l'information en 12 fichiers composés d'enregistrements de taille fixe (1 au niveau Ménage, 2 au niveau Individu, et 9 fichiers biographiques correspondant aux différentes trajectoires biographiques). Cette organisation de l'information satisfait les objectifs de l'enquête relatifs aux analyses aux niveaux individuel comme collectif (ménage, famille) et à l'analyse statistique des données biographiques.

Pour le volet biographique, chaque fichier comporte un enregistrement par étape, incluant les données suivantes : identifiant, date de début, date de fin (seulement pour les trajectoires non continues, c'est-à-dire Éducation, Activité, Nuptialité et Corésidence), ainsi que les variables décrivant éventuellement l'état durant l'étape (par exemple, l'activité ou le lieu de résidence).

La *saisie* des données a été faite sur micro-ordinateur, avec un programme écrit en CLIPPER par le responsable de l'Unité de systèmes du Cede (Argemiro Morales). Écrit spécialement pour cette enquête, et tirant parti de la possibilité offerte par le logiciel de maintenir ouverts simultanément tous les fichiers du questionnaire, le programme de saisie offre les avantages suivants :

— les différents fichiers s'ouvrent successivement de façon automatique, de façon transparente pour l'opérateur. Ainsi, l'opérateur peut suivre la séquence des sections du questionnaire et saisir l'information telle qu'elle apparaît sur le questionnaire, sans une étape préalable de transformation des réponses autre que celle de codification déjà décrite ;

— pendant la saisie, sont réalisés des contrôles de valeurs de toutes les variables, et sont renseignées automatiquement certaines variables en fonction des valeurs prises pour certaines variables de filtre ou des variables particulières.

4.3. Choix de la saisie de la variable temps

Les *variables temps* sont saisies de la façon suivante :

• Section « Étapes migratoires » : elles sont saisies sous leur forme de collecte, c'est-à-dire en nombre d'années vécues dans les différentes catégories de lieu. Les informations figurant dans le tableau, qui servent uniquement à faciliter le calcul de ce nombre d'années en cas de trajectoire complexe, ne sont pas saisies ;

• Section « Système de résidence » : les nombres de jours figurant dans les questions 1 et 3, et le rythme du séjour de la question 4 sont saisis sous la forme décrite précédemment (4.1), soit : nombre de jours et variable synthétique décrivant le rythme du séjour ;

• Section « Biographie » : les dates sont saisies telles qu'elles ont été collectées, c'est-à-dire sous forme d'années, sans autre précision.

• *Volet biographique du deuxième passage à Noida*

La saisie des données du questionnaire de l'enquête renouvelée à Noida est en cours. Pour le volet biographique, les données sont organisées en une série de fichiers biographiques correspondant aux différentes sections de la biographie, en adoptant la structure standard des logiciels spécialisés dans l'analyse de ce type de données, c'est-à-dire un enregistrement par étape comprenant les données suivantes : identifiant, date de début, date de fin éventuelle, variables décrivant l'état durant l'étape (par exemple : lieu de résidence, type de logement, ou emploi). Pour les étapes encore en cours au moment de l'enquête, un code spécifique est assigné au lieu de la date de fin d'étape.

4.3. Choix de la saisie de la variable temps

• *Enquête statistique principale et auprès des sans-logis*

Les variables temps sont saisies telles qu'elles ont été collectées et codées, c'est-à-dire :

Section « Caractéristiques générales » : l'âge est saisi en années révolues, et la date de naissance par l'année sans autre précision.

Section « Étapes migratoires » : les années figurant dans le tableau chronologique d'enregistrement des étapes résidentielles ne sont pas saisies. Les années correspondant aux étapes migratoires clefs sont saisies par leur année (sans autre précision). La durée de séjour dans l'agglomération urbaine de Delhi, et la durée de séjour dans le logement actuel, sont saisies en nombre d'années révolues.

Section « Système de résidence » : la durée de séjour dans chaque lieu est saisie en nombre de jours. La définition des modalités de la variable synthétique décrivant la fréquence de séjour fait également intervenir un décompte du nombre de jours dans chaque lieu de séjour.

Section « Caractéristiques des membres de la famille ne vivant pas dans le logement » : l'âge (actuel ou au décès) est saisi en années révolues, et la date de naissance par l'année sans autre précision. La date du décès est saisie par l'année sans autre précision, et par le nombre d'années (révolues) écoulées depuis le décès.

• *Volet biographique du deuxième passage à Noida*

Les dates du calendrier du volet biographique de l'enquête renouvelée sont saisies telles qu'elles ont été collectées dans la colonne correspondante, c'est-à-dire par leur année sans autre précision.

4.4. Choix de l'unité de temps pour l'analyse

L'unité de temps retenue pour l'analyse varie selon les chapitres du questionnaire, suivant les options retenues pour la codification et la saisie (cf. 4.3) :

biographique. Celui-ci est en cours de réalisation pour le questionnaire définitif à partir d'une programmation propre avec le logiciel Fox Pro.

La « EBIF » comprend cinq bases de données :

1. caractéristiques des logements ;
2. caractéristiques des membres du ménage ;
3. histoire de vie annuelle ;
4. histoire de vie mensuelle pour les cinq dernières années ;
5. antécédents socio-économiques des ascendants et des proches parents.

4.3. Choix de la saisie de la variable temps

La saisie de la variable temps est directement liée à la forme de cette variable dans la collecte. C'est-à-dire que le calendrier des années vécues a été saisi directement, date et année apparaissent comme deux variables, et l'année vécue est l'unité d'analyse. Pour l'histoire mensuelle, une autre base de données a été construite sur le même modèle, le mois vécu étant l'unité d'analyse.

4.5. Utilisation des méthodes d'analyse des biographies

L'analyse des résultats de l'enquête n'a pas été très poussée puisqu'il s'agit d'une enquête pilote (avec un échantillon réduit et non aléatoire) dont l'objectif principal était de tester la méthodologie d'enquête et le questionnaire, en vue de la préparation et de la réalisation d'une enquête représentative au niveau national pour le Mexique (la « EDER », qui est née des résultats de l'expérience « EBIF », a un questionnaire modifié, et a été appliquée en octobre 1998, dans le cadre d'un accord entre l'équipe de chercheurs et l'Inegi mexicain).

Les fonctions de séjour de deux variables ont été étudiées : l'âge à la première migration et l'âge au premier emploi, à partir de l'analyse des tables de séjour et l'estimateur de la fonction de séjour Kaplan-Meier.

4.6. Autres méthodes d'analyse

Un autre type d'analyse a été réalisé pour mesurer le biais de l'échelle annuelle dans la mesure de la migration par rapport à une échelle mensuelle (comparaison des nombres moyens de migrations et durées moyennes de séjour).

4.7. Les utilisateurs des données

Ces données n'ont guère été diffusées au-delà des chercheurs de l'équipe.

4.8. Confrontation avec d'autres enquêtes biographiques

Les résultats n'ont pas été confrontés avec ceux d'autres enquêtes biographiques.